Knowledge Graph Self-Supervised Rationalization for Recommendation

Yuhao Yang
University of Hong Kong
yuhao-yang@outlook.com

Chao Huang*
University of Hong Kong
chaohuang75@gmail.com

Lianghao Xia
University of Hong Kong
aka_xia@foxmail.com

Chunzhen Huang
Wechat, Tencent
chunzhuang@tencent.com

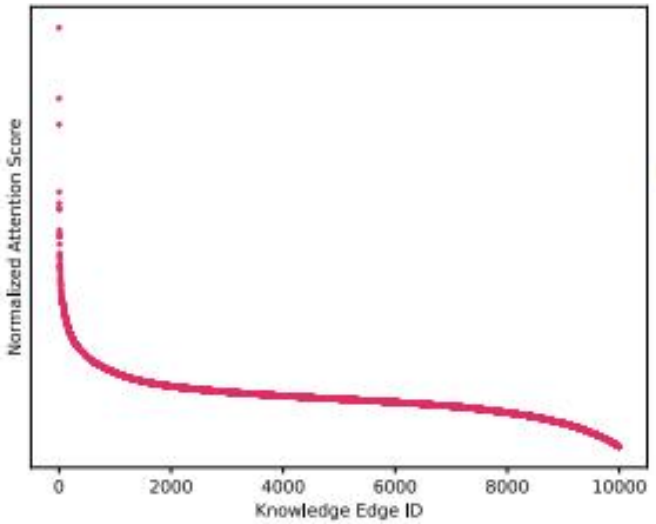
https://github.com/HKUDS/KGRec.

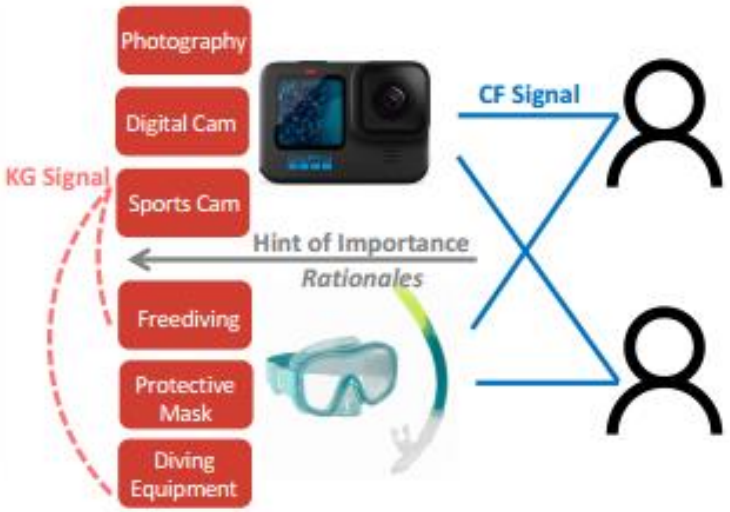
（KDD-2023)

Reported by Ke Gan

1. Introduction
2. Approach
3. Experiments

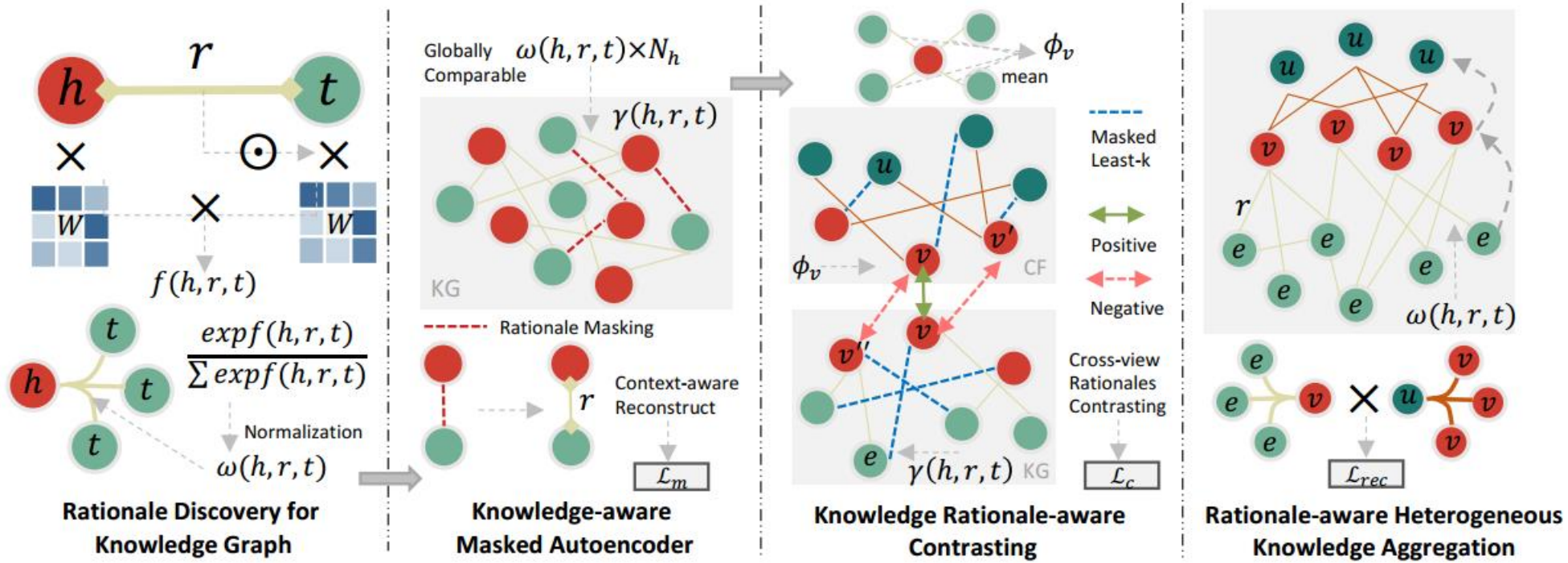
Introduction





adopt either simple random augmentation or intuitive cross-view information, failing to consider the important latent rationales between the KG and recommendation task

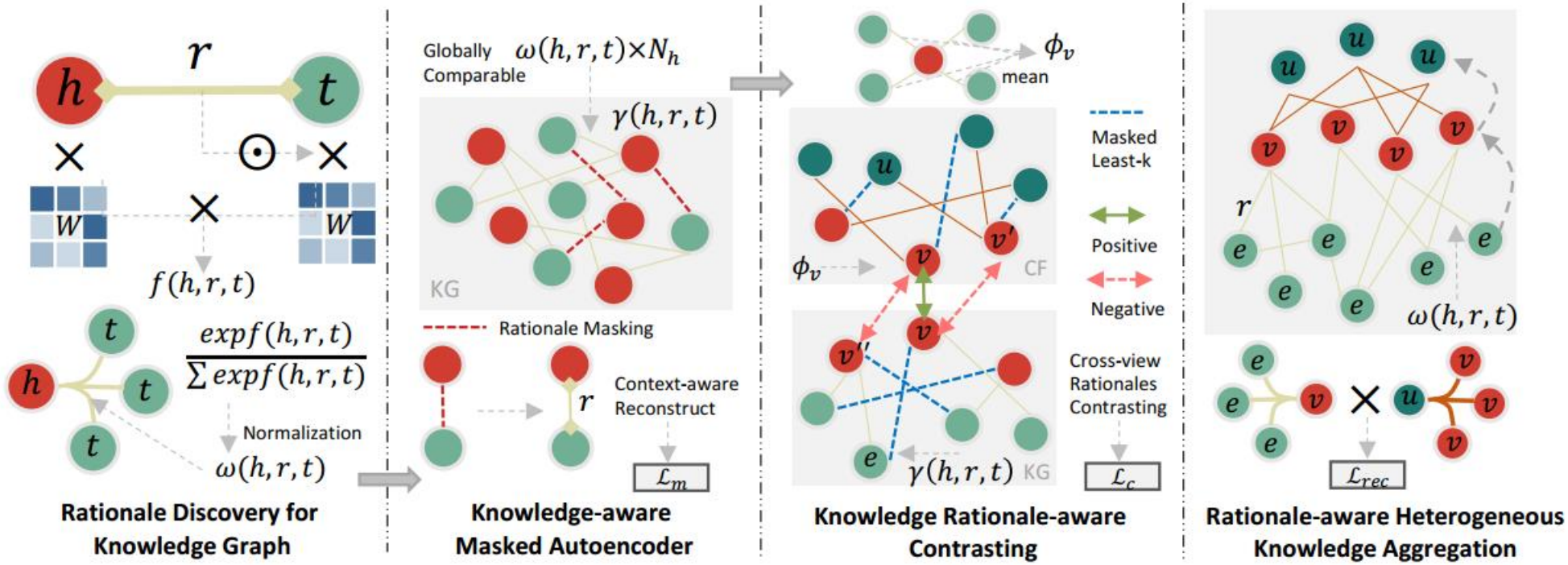
Approach



a binary graph $\mathcal{G}_u = (u, y_{uv}, v)$

$\mathcal{G}_k = (h, r, t).\ h, t \in \mathcal{E}$

Approach



Rational Discovery for KG

$$f(h,r,t) = \frac{\mathbf{e}_h \mathbf{W}^Q \cdot \left(\mathbf{e}_t \mathbf{W}^K \odot \mathbf{e}_r\right)^{\mathsf{T}}}{\sqrt{d}}, \tag{1}$$

$$\omega(h,r,t) = \frac{\exp\left(f(h,r,t)\right)}{\sum_{(h,r',t')\in\mathcal{N}_h} \exp\left(f(h,r',t')\right)}. \tag{2}$$

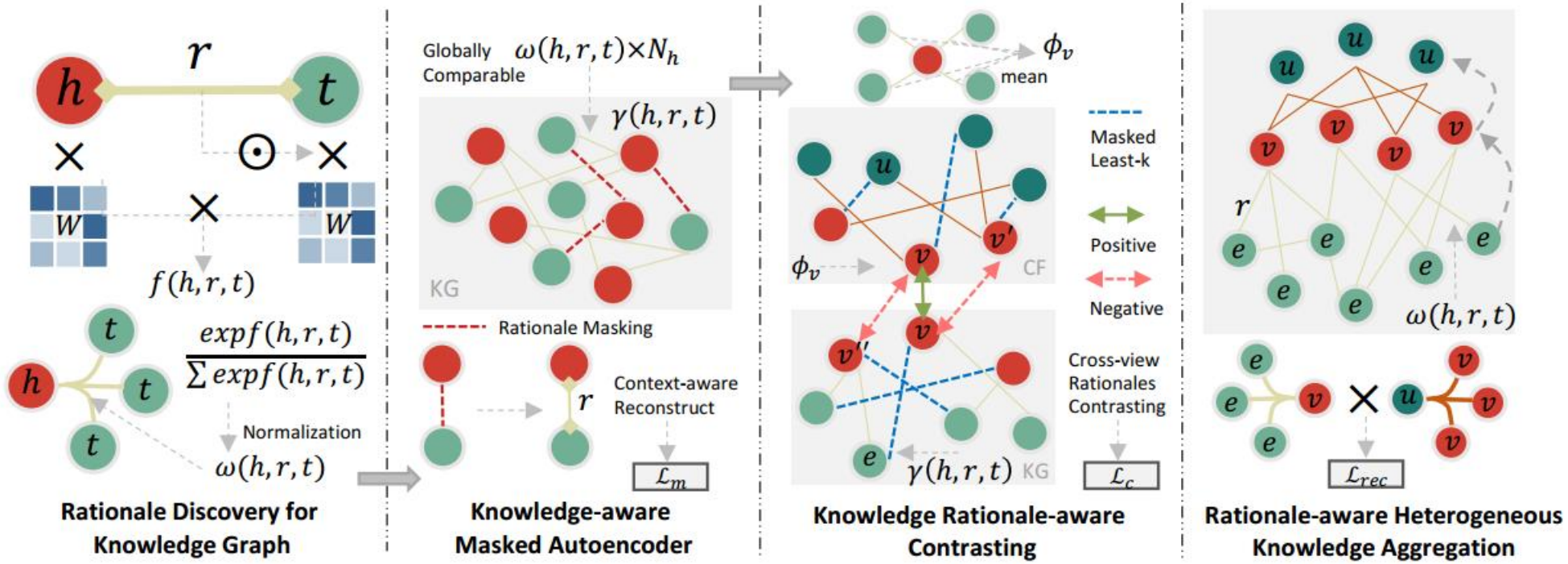
Knowledge Aggregation

$$\mathbf{e}_h^{(l)} = \frac{1}{|\mathcal{N}_h|} \sum_{(h,r,t)\in\mathcal{N}_h} \omega(h,r,t)\mathbf{e}_r \odot \mathbf{e}_t^{(l-1)}, \tag{3}$$

$$\mathbf{e}_u^{(l)} = \frac{1}{|\mathcal{N}_u|} \sum_{i\in\mathcal{N}_u} \mathbf{e}_v^{(l-1)}, \tag{4}$$

$$\mathbf{e}_h = f_k(\mathcal{G}_k; h) = \sum_l^L \mathbf{e}_h^{(l)};\ \mathbf{e}_u = f_u(\mathcal{G}_u; u) = \sum_l^L \mathbf{e}_u^{(l)}, \tag{5}$$

Approach



Knowledge-aware Masked Autoencoder

$$\gamma(h,r,t) = |\mathcal{N}_h| \cdot \omega(h,r,t) = \frac{|\mathcal{N}_h| \cdot \exp\left(f(h,r,t)\right)}{\sum_{(h,r',t')\in\mathcal{N}_h} \exp\left(f\ (h,r',t')\right)}. \quad (6)$$

$$\gamma(h,r,t) = \gamma(h,r,t) - \log\left(-\log(\epsilon)\right); \quad \epsilon \sim \text{Uniform}\ (0,1), \quad (7)$$

$$\mathcal{M}_k = \{(h,r,t)|\gamma(h,r,t) \in \text{topk}(\Gamma; k_m)\}, \quad (8)$$

$$\mathbf{e}_h = f_k(\mathcal{G}_k^m; h);\ \mathbf{e}_t = f_k(\mathcal{G}_k^m; t), \quad (9)$$

$$\mathcal{L}_m = \sum_{(h,r,t)\in\mathcal{M}_k} -\log\left(\sigma\left(\mathbf{e}_h^{\mathsf{T}} \cdot (\mathbf{e}_t \odot \mathbf{e}_r)\right)\right). \quad (10)$$

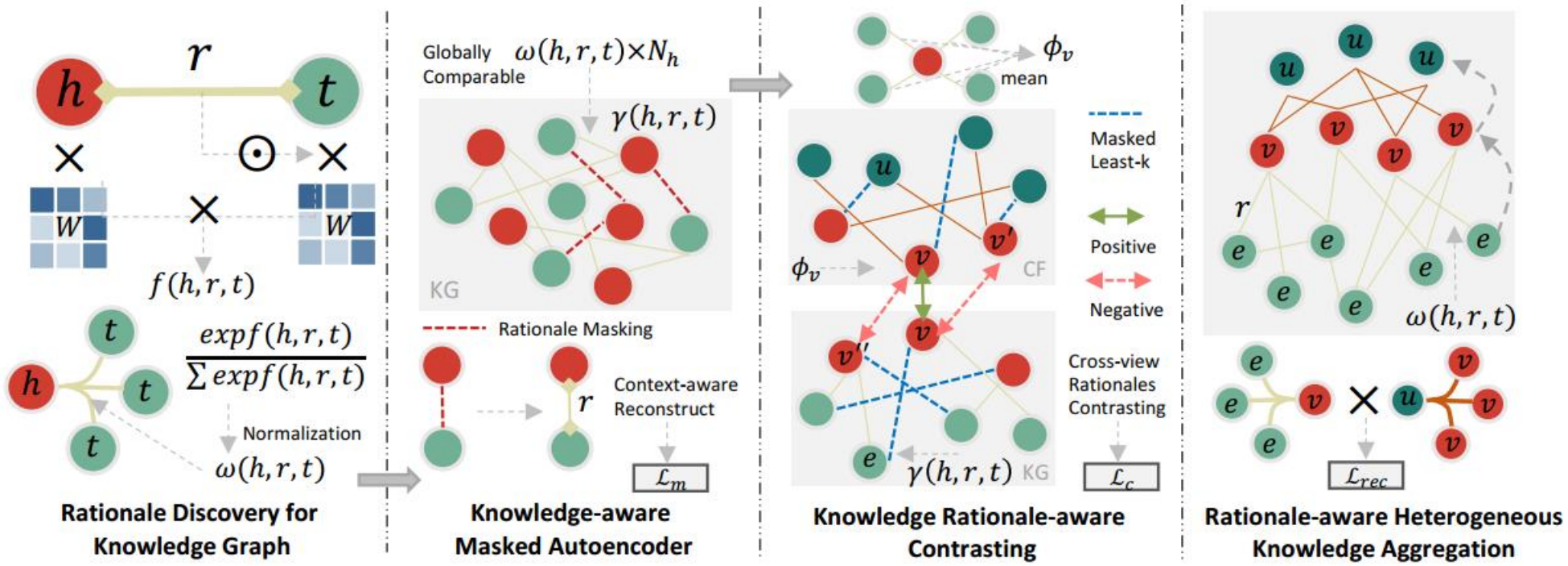
Knowledge Rationale-aware Contrasting

$$\mathcal{S}_k = \{(h,r,t)|\gamma(h,r,t) \in \text{topk}(-\Gamma; \rho_k)\};\ \mathcal{G}_k^c = \mathcal{G}_k \setminus \mathcal{S}_k, \quad (11)$$

$$\phi_v = \text{mean}(\{\gamma(h,r,t)|h = v \vee t = v\}). \quad (12)$$

$$\phi_v' = \frac{\exp\phi_v}{\sum_v \exp\phi_v};\ \mathcal{S}_u \sim \text{multinomialNR}(\Phi'; \rho_u), \quad (13)$$

Approach



$$\mathbf{x}_u^{(l)} = \sum_{v \in \mathcal{N}_u} \frac{\mathbf{x}_v^{(l-1)}}{\sqrt{|\mathcal{N}_u||\mathcal{N}_v|}}; \ \mathbf{x}_v^{(l)} = \sum_{u \in \mathcal{N}_v} \frac{\mathbf{x}_u^{(l-1)}}{\sqrt{|\mathcal{N}_u||\mathcal{N}_v|}}. \tag{14}$$

$$\mathbf{x}_v^k = f_k(\mathcal{G}_k^c; v). \tag{15}$$

$$\mathbf{z}_v^* = \sigma\left(\mathbf{x}_v^{*\mathsf{T}}\mathbf{W}_1^* + \mathbf{b}_1^*\right)^{\mathsf{T}}\mathbf{W}_2^* + \mathbf{b}_2^*, \tag{16}$$

$$\mathcal{L}_c = \sum_{v \in \mathcal{V}} -\log \frac{\exp(s(\mathbf{z}_v^u, \mathbf{z}_v^k)/\tau)}{\sum_{j \in \{v, v', v''\}} (\exp(s(\mathbf{z}_v^u, \mathbf{z}_v^k)/\tau) + \exp(s(\mathbf{z}_j^u, \mathbf{z}_v^k)/\tau))}, \tag{17}$$

$$\mathcal{L}_{rec} = \sum_{(u,v,j) \in \mathcal{D}} -\log \sigma\left(\hat{y}_{uv} - \hat{y}_{uj}\right), \tag{18}$$

$$\mathcal{L} = \mathcal{L}_{rec} + \lambda_1 \mathcal{L}_m + \lambda_2 \mathcal{L}_c, \tag{19}$$

Experiment

Table 1: Statistics of Three Evaluation Datasets.

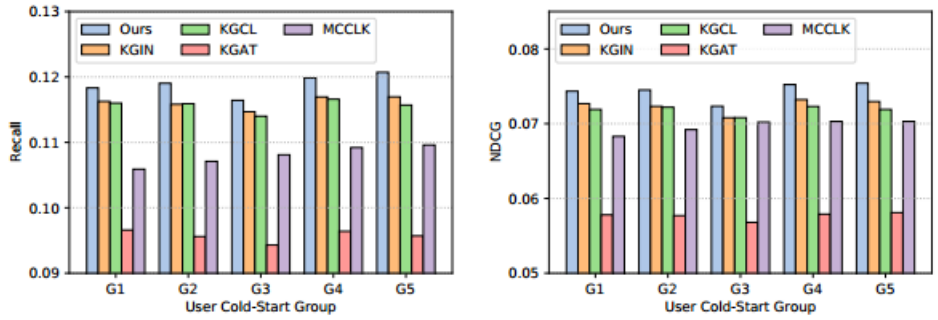
Statistics	Last-FM	MIND	Alibaba-iFashion
# Users	23,566	100,000	114,737
# Items	48,123	30,577	30,040
# Interactions	3,034,796	2,975,319	1,781,093
# Density	2.7e-3	9.7e-4	5.2e-4
Knowledge Graph			
# Entities	58,266	24,733	59,156
# Relations	9	512	51
# Triplets	464,567	148,568	279,155

Model	Last-FM		MIND		Alibaba-iFashion	
	Recall	NDCG	Recall	NDCG	Recall	NDCG
BPR	0.0690	0.0585	0.0384	0.0253	0.0822	0.0501
NeuMF	0.0699	0.0615	0.0308	0.0237	0.0506	0.0276
GC-MC	0.0709	0.0631	0.0386	0.0261	0.0845	0.0502
LightGCN	0.0738	0.0647	0.0419	0.0253	0.1058	0.0652
SGL	0.0879	0.0775	0.0429	0.0275	0.1141	0.0713
CKE	0.0845	0.0718	0.0387	0.0247	0.0835	0.0512
KTUP	0.0865	0.0671	0.0362	0.0302	0.0976	0.0634
KGNN-LS	0.0881	0.0690	0.0395	0.0302	0.0983	0.0633
KGCN	0.0879	0.0694	0.0396	0.0302	0.0983	0.0633
KGAT	0.0870	0.0743	0.0340	0.0287	0.0957	0.0577
KGIN	0.0900	0.0779	0.0357	0.0225	0.1144	0.0723
MCCLK	0.0671	0.0603	0.0327	0.0194	0.1089	0.0707
KGCL	0.0905	0.0769	0.0399	0.0247	0.1146	0.0719
KGRec	**0.0943**	**0.0810**	**0.0439**	**0.0319**	**0.1188**	**0.0743**

Experiment

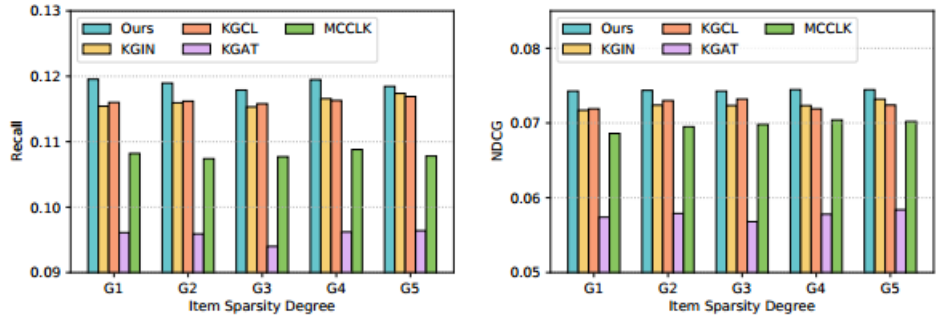
Table 3: Ablation results of KGRec with different variants. The superscript * denotes the largest change in performance.

Ablation Settings	Last-FM		MIND		Alibaba-iFashion	
	Recall	NDCG	Recall	NDCG	Recall	NDCG
KGRec	**0.0943**	**0.0810**	**0.0439**	**0.0319**	**0.1188**	**0.0743**
w/o MAE	0.0918*	0.0792*	0.0374*	0.0238*	0.1178*	0.0737*
w/o Rationale-M	0.0929	0.0805	0.0423	0.0311	0.1183	0.0739
w/o CL	0.0926	0.0796	0.0425	0.0313	0.1180	0.0734
w/o Rationale-Aug	0.0931	0.0801	0.0405	0.0278	0.1185	0.0741



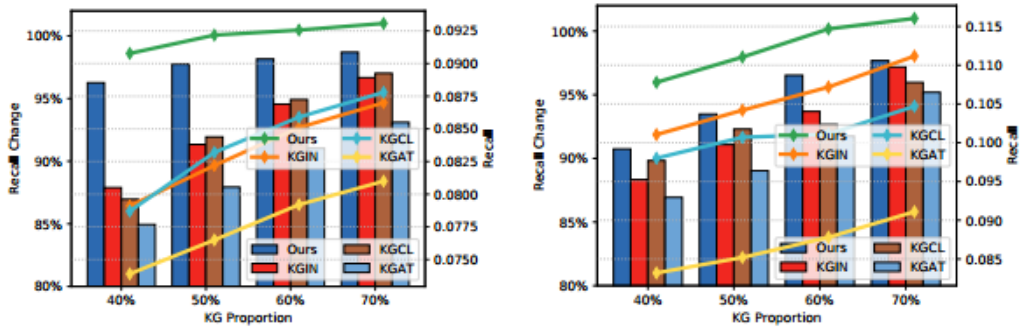
(a) Alibaba-iFashion@Recall (b) Alibaba-iFashion@NDCG

Figure 3: Evaluation results on different user groups. Lower group number implies stronger cold-start effect.



(a) Alibaba-iFashion@Recall (b) Alibaba-iFashion@NDCG

Figure 4: Evaluation results on different item sparsity levels.



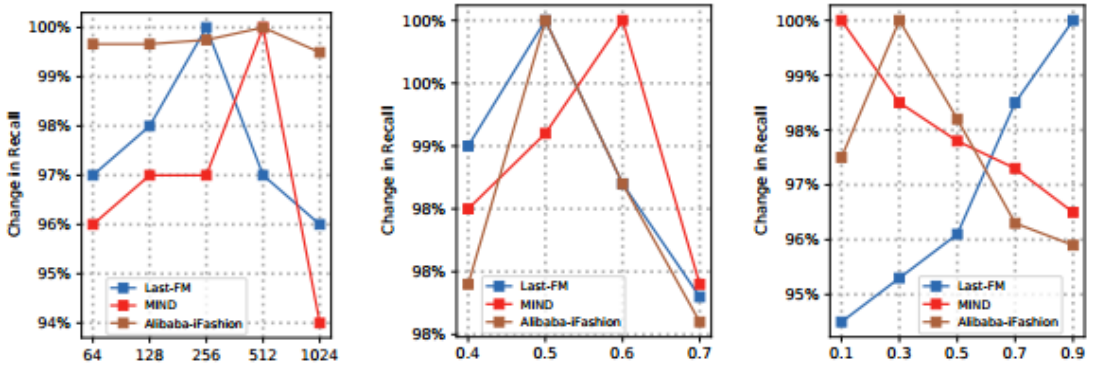
(a) Last-FM (b) Alibaba-iFashion

Figure 5: Evaluation results on different KG proportions.

Experiment

Table 4: KG Relations with highest average global rationale scores for news categories learned on MIND dataset.

Category	Relation (Wiki ID)	Avg. Score
sports	member of sports team (P54)	1.235
	league of (P118)	1.117
newspolitics	member of political party (P102)	1.341
	position held (P39)	1.097
travel	part of (P361)	1.105
	located in (P131)	1.190
finance	owned of (P1830)	1.203
	stock exchange (P414)	1.157
tv-celebrity	award received (P166)	1.084
	cast member (P161)	1.139



(a) Masking size k_m (b) CL keep ratio ρ (c) CL temperature τ

Figure 6: Hyperparameter Study of KGRec.

Thank you!